It's a Jungle Out There: Fantasy and Reality of Evaluating Public Displays "in the Wild"

Timo Ojala
University of Oulu
P.O.Box 4500
90014 University of Oulu, Finland
timo.ojala@ee.oulu.fi

Vassilis Kostakos
Madeira Interactive Technologies Institute
University of Madeira
9020-105 Funchal, Portugal
vk@uma.pt

Hannu Kukka
University of Oulu
P.O.Box 4500
90014 University of Oulu, Finland
hannu.kukka@ee.oulu.fi

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

Abstract
We present a summary of our experiences with a longitudinal deployment of interactive public displays in a city center. We argue that such deployments offer external validity and highlight a number of issues overlooked by lab studies. We also point out that these benefits come with substantial cost and difficulties.

Keywords
Interactive public displays, longitudinal evaluation

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction
A common criticism targeted at many studies on interactive public displays is that their evaluation usually takes place in non-realistic lab environments, and for short periods of time. Thus, a long-term real-world deployment could be a more appropriate evaluation. However, while it promises high external validity, it also entails a series of obstacles, difficulties, and constraints. This paper summarizes successes and difficulties from a deployment of 12 public displays in a city center, having lasted soon two years, during which thousands of real users have used the displays.

Fantasy

In our long-term vision the urban space is populated with hotspots, which provide rich interaction between the physical, virtual and social spaces. We do not offer these hotspots as solutions to particular (research) problems in the urban space, which undoubtedly would survive just fine without them. Instead, we promote the hotspots as 'heavyweight' urban probes [3], instruments for measuring an unknown and producing hopefully useful and interesting data. In our case the unknown is the urban landscape, where new technology is emerging and where complex social roles of the urban communities, people's movement and traces through cities, and people's interactions with place and public artifacts intersect. By assessing the reaction and interaction of people and place with the probe we hope to learn more about the urban landscape.

We have realized our vision with a concept dubbed 'UBI-hotspot' (Fig. 1). It is effectively a large interactive public display embedded with other co-located computing resources such as cameras and wireless access points. The hotspots are versatile computing platforms, which facilitate provisioning of a diverse range of services to the general public in authentic urban setting. The hotspots enable in-situ urban computing research with real users and with sufficient scale and time span. Such studies are important because real world systems are culturally situated, and cannot be reliably assessed with lab studies detached from the real world context. By deploying a number of hotspots for a sufficiently long time, we wish to establish technical and cultural readiness, and the critical mass of users, needed for determining whether our concept can be deemed '(un)successful' [1].

Reality

System description

We deployed a network of 12 hotspots at pivotal outdoor and indoor locations around downtown Oulu in summer 2009. The six double-sided outdoor hotspots are installed at the walking street area at the heart of the city and at the market area. The six single-sided indoor hotspots are placed in popular municipal buildings such as main library, youth and culture center, and swimming hall.

Figure 1. Outdoor UBI-hotspot at downtown Oulu.

An idle hotspot is in a passive broadcast mode where the whole screen shows so-called UBI-channel, a customizable playlist of video, animation and still photographs. The hotspot changes to an interactive mode, if a face is detected by the overhead cameras or the screen is touched. In the interactive mode the screen is split between the UBI-channel, and an interactive portal dubbed UBI-portal. It comprises of various information and leisure web services, which are referenced by their URLs and can reside on any web server in the public Internet. So far, we have released two versions of the portal, version 1 in June 2009 [2]

and version 2 in June 2010. Version 1 attracted on average 1564 clicks per day during a 42-week observation period in Jul 2009 - Apr 2010. Version 2 attracted on average 948 clicks per day during 41 weeks in Jun 2010 - Feb 2011.

Differences between the lab and the wild

The first important conclusion we have arrived at is that there exists a huge difference between results obtained in a lab and in the wild using the exact same configuration. While in lab tests participants were quite happy to interact with a hotspot and explore its functionality, it has been challenging to entice the general public to interact with our real-world hotspots. Further, although several usability experts both from industry and academia have been involved in the design and validation of the UI, we have observed that general public found it hard to navigate and use the hotspot. These findings highlight important real world issues often ignored in lab studies: would real users have the motivation and the skills to use a particular public display? Generally speaking, how representative are user samplings in lab studies?

Another important finding is that location is absolutely crucial to the way our systems are used by general public. Even though we have simultaneously deployed identical configurations at multiple locations, we have observed striking differences at how the hotspots are used and which services are most popular. For example, a hotspot placed in the lobby of a swimming hall attracted 47 times more clicks than an identical hotspot placed in the lobby of a municipal service center [2]. The swimming hall with patrons in relaxed leisurely mood, especially lots of children and teens keen to play games, proved to be much more suitable location that the business like, almost clinical, municipal service center. Hence, we argue that the effects of 'location' are crucial dimensions that lab studies or one-shot studies cannot effectively capture.

A further crucial aspect of the usage of our hotspots is the effect of curiosity, which is difficult to replicate in a lab study. We have observed a direct effect on the number of clicks required the access a service and its usage numbers. The effect is not linear, however, as some services attract a clear user base who access the service despite requiring multiple clicks and despite the service not being promoted in the opening 'quick launch' menu. With other services the effect is more pronounced, with usage dropping to near zero once the service is taken away from the 'quick launch' menu, even though it had fair usage while in the menu. This leads us to hypothesize that a substantial amount of the usage we observe on our hotspots is attributed to curiosity. We argue that passersby are curious about something on the screen, and are willing to make a single click to observe the response of the system. One suggested mechanism for weeding out curiosity usage is by incorporating small barriers to usage (such as the need to have a short but meaningful interaction with the system) before the system can actually be used. We expect that the incorporation of such a mechanism will reduce overall usage of the hotspots, but at the same time we expect the reduction to be mostly attributed to curious users being omitted.

We have also observed a systematic impact of novelty: when new features are introduced in the system or when a major system upgrade takes place, usage first increases, but then gradually decreases. Further, weather appears to have an effect so that sunnier and

warmer days correlate with higher usage. These two effects, novelty and weather, are difficult to replicate in a lab study.

In summary, we have identified crucial differences between the lab and the wild, which we attribute to user sampling, location, curiosity, novelty, and weather. These substantial differences lead us to question the validity of lab studies in this domain.

Challenges

While we feel that our deployment has given us a unique perspective on interactive public displays, we have faced many obstacles. The greatest challenge in our deployment has been maintenance. A substantial amount of our resources has been used to ensuring that the system behaves as expected, to troubleshoot faults, and even to clean the displays. Vandalism is another issue we have had to deal with.

We have invested significant effort in ensuring that the community feels engaged with our deployment. During summer months we have hosted weekly 'UBI walks' that general public can attend to learn to use the hotspots. The City of Oulu is allocated with a significant portion of the capacity of the UBI-channel to advertise public services and events, running 212 distinct campaigns in 2010. We have also organized several competitions to motivate the local community to contribute ideas for services on the hotspots.

Finally, the deployment is challenged by economic viability and measurement of success. First, after the initial 400 000 € capital investment it has been challenging to ensure that the hotspots become self-sufficient. We have generated revenue from commercial use, but this has conflicted with research objectives. When we sell a given capacity of the UBI-channel, it has to be visible which constrains screen layout and interaction model. Second, the discipline as a whole lacks standardized metrics for evaluating the success of such systems. In other words, despite the availability of huge volumes of quantitative and qualitative data on the usage of the hotspots, it is challenging to measure their actual impact on the community.

Conclusion

This paper argues that real-world longitudinal deployment of interactive public displays reveals a number of factors affecting their use that lab studies do not necessarily take into account: user sampling, location, curiosity, novelty, and weather. We also argue that such deployments entail a number of challenges, including maintenance, vandalism, engaging the local community, sustainability, and measurement of impact.

Acknowledgments

The authors wish to thank the Finnish Funding Agency for Technology and Innovation, the Academy of Finland, the ERDF, the City of Oulu and the UBI (UrBan Interactions) consortium for their valuable support.

References

[1] Greenberg, S. and Buxton, B. Usability evaluation considered harmful (some of the time). *Proc. CHI 2008*, 111-120.

[2] Ojala, T., Kukka, H., Lindén, T., Heikkinen, T., Jurmu, M., Hosio, S. and Kruger, F. UBI-hotspot 1.0: Large-scale long-term deployment of interactive public displays in a city center. Proc. *ICIW 2010*, 285-294.

[3] Paulos, E. and Jenkins, T. Urban Probes: Encountering our emerging urban atmospheres. *Proc. CHI 2005*, 341-350.